Expectation and variation with a virtual die

Jane Watson
University of Tasmania
Jane.Watson@utas.edu.au

Lyn English
Queensland University of Technology
l.english@qut.edu.au

Introduction

By the time students reach the middle years they have experienced many chance activities based on dice. Common among these are rolling one die to explore the relationship of frequency and theoretical probability, and rolling two dice and summing the outcomes to consider their probabilities. Although dice may be considered overused by some, the advantage they offer is a familiar context within which to explore much more complex concepts. If the basic chance mechanism of the device is understood, it is possible to enter quickly into an arena of more complex concepts. This is what happened with a two hour activity engaged in by four classes of Grade 6 students in the same school.

The activity targeted the concepts of variation and expectation. The teachers held extended discussions with their classes on variation and expectation at the beginning of the activity, with students contributing examples of the two concepts from their own experience. These notions are quite sophisticated for Grade 6, but the underlying concepts describe phenomena that students encounter every day. For example, time varies continuously; sporting results vary from game to game; the maximum temperature varies from day to day. However, there is an expectation about tomorrow's maximum temperature based on the expert advice from the weather bureau. There may also be an expectation about a sporting result based on the participants' previous results. It is this juxtaposition that makes life interesting. Variation hence describes the differences we see in phenomena around us. In a scenario displaying variation, expectation describes the effort to characterise or summarise the variation and perhaps make a prediction about the message arising from the scenario. The explicit purpose of the activity described here was to use the familiar scenario of rolling a die to expose these two concepts.

Because the students had previously experienced rolling physical dice they knew instinctively about the variation that occurs across many rolls and about the theoretical expectation that each side should "come up" one-sixth of the time. They had observed the instances of the concepts in action, but had not consolidated the underlying terminology to describe it. As the two concepts are so fundamental to understanding statistics, we felt it would be useful to begin building in the familiar environment of rolling a die. Because hand-held dice limit the explorations students can undertake, the classes used the soft-ware *TinkerPlots* (Konold & Miller, 2011) to simulate rolling a die multiple times.

What outcomes can we expect?

The activity included a discussion of students' previous experiences of rolling dice, whether some numbers were more likely to come up than others and whether students had "lucky" numbers. When asked if one side were more likely to come up, only a few students recorded a response on their worksheets, with reasons ranging from "5 because 1 and 6 are harder to get," to, "I mostly get 2," and, "3 because it is in the centre."

The teachers reviewed the students' thinking about the chances of outcomes when rolling a die by confirming the theoretical expectation (probability) of $\frac{1}{6}$, where each side is equally likely to come up. They then asked students to speculate on the expected outcomes of rolling a die 30 times. In stating their predictions on their worksheets, most students suggested that each outcome would have 5 instances. Of those who did not, some had difficulty allocating the total of 30 outcomes among the six possibilities. Students' reasons included stating that "random" outcomes should not be even, telling stories of students' personal experiences, expressing beliefs about lucky numbers, or expecting clusters of results around the middle of the numbers. Typical of the explanations for the response of 5 for each side of the die was, "I think this because they all have an equal chance of being rolled and 6 × 5 = 30." Near the start of the activity the teachers explicitly discussed the translation of the fraction 1/6 into a percentage, rounded to 17%. The equivalence of fractions, decimals, and percentages is an expectation of Year 6 of the *Australian Curriculum: Mathematics* (ACARA, 2013) and all students demonstrated on the worksheet that they could do this.

Investigations with a virtual die: Four steps to making decisions with data

The activity was implemented through two investigations formulated to reinforce the investigative steps in Figure 1, Four Steps to Making Decisions with Data, and using *TinkerPlots* to carry out the simulations.

First investigation

The first investigation was based on the question (Step 1) about which students had speculated: "What do you expect the outcomes will be when you roll a die 30 times?" The worksheets included illustrated instructions for setting up the Sampler (a simulator) in *TinkerPlots* to carry out the simulations of the 30 rolls (see Appendix).

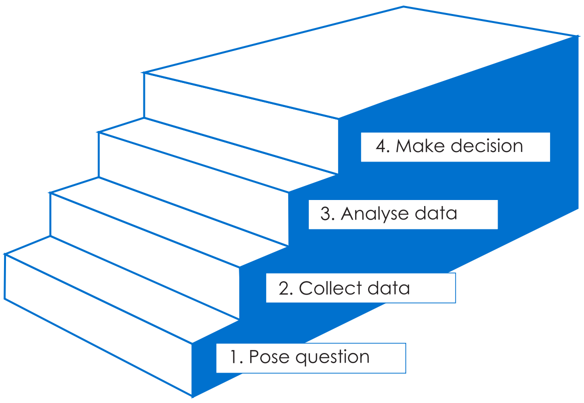


Figure 1. Four steps to making decisions with data.

It was felt that constructing the simulation process in the Sampler would add to students' appreciation of how the software was imitating the procedure they might follow if they actually rolled a die themselves. Each step in the investigation was highlighted on the worksheet with a smaller icon than shown in Figure 1. Steps 2 and 3 were completed by students with *TinkerPlots* as shown in the Appendix, with data collection (Step 2) carried out by the Sampler and the data analysis (Step 3) based on the Plot (a graphical representation). All work in *TinkerPlots* was set up by the students, working in pairs.

With output from *TinkerPlots* that looked like the format in Figure 2, students were asked to record the largest and smallest percentage values and find the difference, that is, the range of values (continuing Step 3). Students had met the range in an earlier activity and the

objective was for students to observe the range decrease as the number of rolls increased. For example, the range in Figure 2 would have been recorded as:

(Largest % value) – (Smallest % value) = (Range of %)

37% – 7% = 30%.

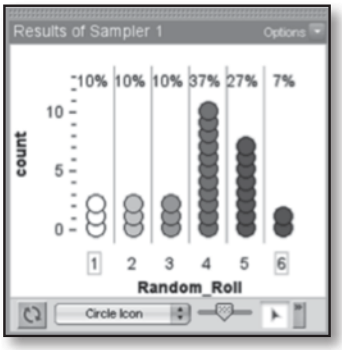


Figure 2. Example of output for 30 rolls of a die from the software *TinkerPlots*.

Students then repeated the simulation with the Sampler four more times, recording the five outcomes in a table. Having the visual representations such as the one in Figure 2 reinforced the large range in percentage for many of the simulations. Based on these values students were asked to (a) decide if their outcomes were close to their expectations, (b) to record a conclusion based on their data (Step 4), (c) to state how confident they were in the conclusion, and (d) to say what would make them more confident. Due to the sharing of their results during the trials, the students had observed much variation in the outcomes. Sometimes their conclusions were expressed in terms of variation but denying any expectation. Two students, for example, wrote:

"It's just random as I can't tell you" and "The range and values were different. It means that you can't expect anything because it is random".

Other students appreciated the underlying expectation as well as the variation. Two students wrote: "It could be any number but it's normally near 5" and "Based on the data I collected I found even though each number has an equal chance the results vary a lot".

These students were quite confident of their conclusions and for most students increased confidence would be based on, "if we did more trials." Some who said, "just random" said nothing would increase their confidence in this conclusion.

Second investigation

The discussion of the large variation in the ranges when 30 rolls were simulated led to the suggestion of the second investigation, which was based on the question (Step 1): "What happens to the "Range of %" as you increase the number of rolls?" Students were provided with tables to record the outcomes of 5 trials of 100 rolls, 5 trials of 1000 rolls, and 5 trials of 10,000 rolls, in each case recording the "Largest %," the "Smallest %," and the "Range of %" for each trial (Step 2). The data from one such record is shown in Figure 3.

1. Change the Repeat number to 100.

	Largest %	Smallest %	Range of %
Trial 1 – 100 rolls	25%	9%	16%
Trial 2 – 100 rolls	25%	13%	12%
Trial 3 – 100 rolls	20%	11%	9%
Trial 4 – 100 rolls	22%	13%	9%
Trial 5 – 100 rolls	23%	9%	14%

2. **Change the Repeat number to 1000.**

	Largest %	Smallest %	Range of %
Trial 1 –1000 rolls	19%	15%	4%
Trial 2 –1000 rolls	18%	15%	3%
Trial 3 –1000 rolls	18%	15%	3%
Trial 4 –1000 rolls	18%	15%	3%
Trial 5 –1000 rolls	18%	16%	2%

3. **Change the Repeat number to 10 000.**

	Largest %	Smallest %	Range of %
Trial 1 –10 000 rolls	17%	16%	1%
Trial 2 –10 000 rolls	17%	16%	1%
Trial 3 –10 000 rolls	18%	16%	2%
Trial 4 –10 000 rolls	17%	16%	1%
Trial 5 –10 000 rolls	17%	16%	1%

Figure 3. Tables for increasing numbers of rolls (Step 2, Data collection from Student A).

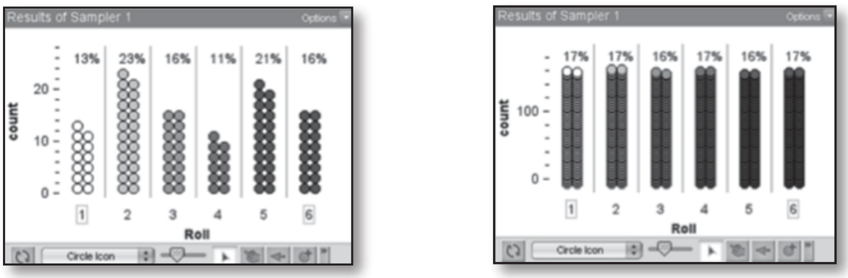
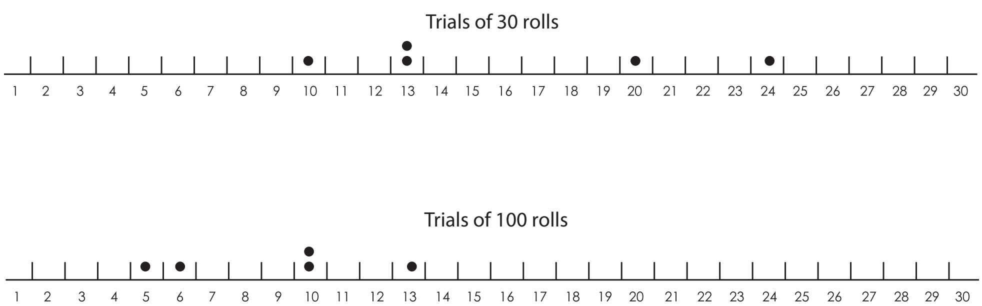


Figure 4. Difference in visual appearance between 100 rolls (left) and 1000 roll (right) of the die.

Students recorded data from looking at more plots as in Figure 2 and hence they had the visual impression of the percentages being closer together as the sample size increased. This is shown in Figure 4 for samples of 100 and 1000 rolls. To analyse the data (Step 3), and further reinforce the reduction in the “Range of %” with increasing number of rolls, students plotted the ranges on four number lines, one line each for the five trials of 30 rolls, 100 rolls, 1000 rolls, and 10 000 rolls. A scan of one sequence of plots from a different student is shown in Figure 5. All of the students concluded (Step 4) that the “Range of %” decreased as the number of rolls increased. Throughout the investigation teachers reminded students of the steps in Figure 1.



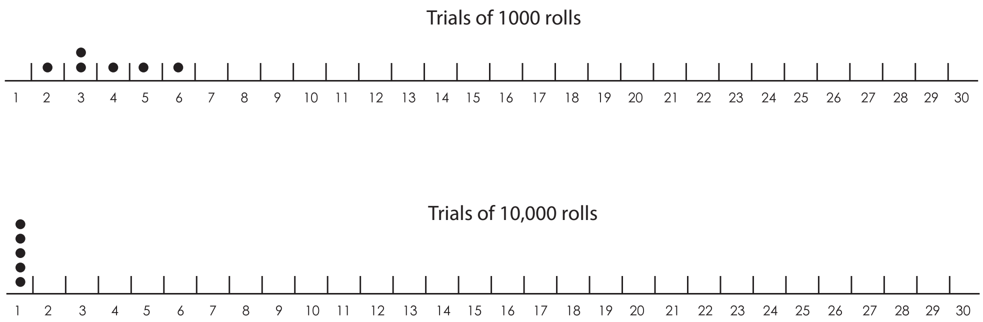


Figure 5. Plots of ranges for increasing numbers of rolls (Step 3, Data analysis from Student B).

What did the students learn?

To assess the students' understanding further, they were asked to write a sentence or two in their workbooks explaining what they had learned about variation and expectation from the activity. To reinforce the language associated with the concepts, they were also asked to use the words "variation" and "expectation" in their sentences. The transition from observing the phenomena to using the words in sentences was difficult for some students, particularly those with English as a second language. In relation to the word expectation, some referred only to their own expectation but did not link to the expectation for the die:

- "I learned that my expectation is not always correct."
- "My expectation of what percentage was going to happen was very surprising!"

Others were able to link expectation more closely to the context:

- "My expectation was all equal outcome[s] from each number."
- "The word expectation means what you expect and we did some 'expectations' by predicting the fractions and percentages of the chance of rolling a 6 sided dice."
- "I have learnt from this activity that the accuracy of my expectation depends on the amount [sic] of trials conducted."

Most students could use variation in a meaningful sentence. Some, however, did not link the word to their learning:

- "I learnt that the answer can vary sometimes in big ways sometimes in small ways."
- "I have learned that variations always change."

Others described variation generally in terms of the activity:

- "The variation of the times the dice rolled was many different numbers."
- "I learned that there is a variation of numbers that can happen."

Further, some students described the effect of increasing the number of rolls:

- "The variation got lower as we increased the number of rolls."
- "I lear[n]t that when you increase the number of trials the variation of % becomes smaller. 30 rolls = 27% – 7% 10 000 = 1% – 1%."

The most sophisticated responses could create a concise description of the activity explicitly juxtaposing variation and expectation. Most of these, but not all, considered increasing numbers of trials.

- "In this activity I learned that when making expectations about rolling a six-sided dice the numbers should be equal and when the numbers of times rolled increase the variation in the ranges will decrease."
- "During this activity, I learnt that our expectation 17% will vary as we increased the number of rolls. As the number of rolls increased the range of % decreased and the outcomes were close."

- “I learned in this lesson that an expectation has a variation but there will always be a common expectation. A variation could be decreased when number of rolls are increased.”
- “From this activity, I have learned that expectation and results have a variation. I have also learned that the more number of trials conducted, the less variation between results and expectation.”
- “My expectation was that if you do more trials you get a similar percentage. The variation of the small[er] the value [of number of trials] the bigger the range.”

Consolidating expectation and variation

Although the students in these classes all appreciated the convergence of the simulated outcomes to the expected probabilities and could describe it colloquially, it was difficult for some to articulate concise appropriate descriptions of their experience using the written language of expectation and variation. The teachers used and reinforced the language throughout the investigations but many different elaborations are required for students to consolidate its use. Specifically, this consolidation can happen during collecting and observing data: “What are we expecting to see?” “How do we describe what we see? What words can we use?” When summarising results, students can be asked, “What did we expect to happen and what can we expect to happen if we do the trials again?” “What is likely to reduce the variation we saw in our results?” The examples of students’ outcomes and descriptions of what they learned about variation and expectation presented here should help teachers to anticipate what might occur in their own classrooms and to plan accordingly.

Other experiences can focus, for example, on taking random samples of increasing size from a known population of measurements and watching the means of the samples approach the mean of the population (For example, Watson, 2006, pp. 242–244). In later years, this activity with a die provides a foundation for introducing the Law of Large Numbers. The law is discussed and illustrated for high school students by Flores (2014) based on games using coins and by Hoffman and Snapp (2012) using dice.

Concluding points

It is hoped that this description of an activity aimed at establishing both the concepts and language of variation and expectation will encourage other middle and high school teachers to conduct similar lessons. Although the *Australian Curriculum: Mathematics* (ACARA, 2013) does not explicitly address increasing sample size in relation to variation and expectation within Statistics and Probability, it can be inferred from the Year 8 descriptor: “Explore the variation of means and proportions of random samples drawn from the same population” (ACMSP293), and its elaboration, “using sample properties to predict characteristics of the population” (p. 54). Sample size is definitely a property of a sample. There is no reason why these ideas should not be introduced and reinforced across the middle years.

Other software and spreadsheets can be used instead of *TinkerPlots* to simulate the rolling of the die. It is likely, however, that the technicalities of setting up the simulation, because of their complexity, would need to be completed by the teacher rather than by Grade 6 students. As seen in the Appendix, *TinkerPlots* is purpose-built software that visually supports the creation of objects and their application. Other instances of the affordances for learning *TinkerPlots* are found in Hudson (2012) supporting understanding

of the mean and in Watson, Fitzallen, Wilson, and Creed (2008) interpreting graphical representations using hat plots.

References

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2013). *The Australian Curriculum: Mathematics* (version 5.0). Sydney, NSW: ACARA. Retrieved 20 May 2013 from http://www.australiancurriculum.edu.au/mathematics/curriculum/f-10?layout=1#cdcode=ACMSP293&level=8

Flores, A. (2014). Who's in the Lead? *The Mathematics Teacher, 108*, 18–22.

Hoffman, T.R. & Snapp, B. (2012). Gaming the Law of Large Numbers. *The Mathematics Teacher, 106*, 378–383.

Hudson, R.A. (2012). Finding the balance at the elusive mean. *Mathematics Teaching in the Middle School, 18*, 301–306.

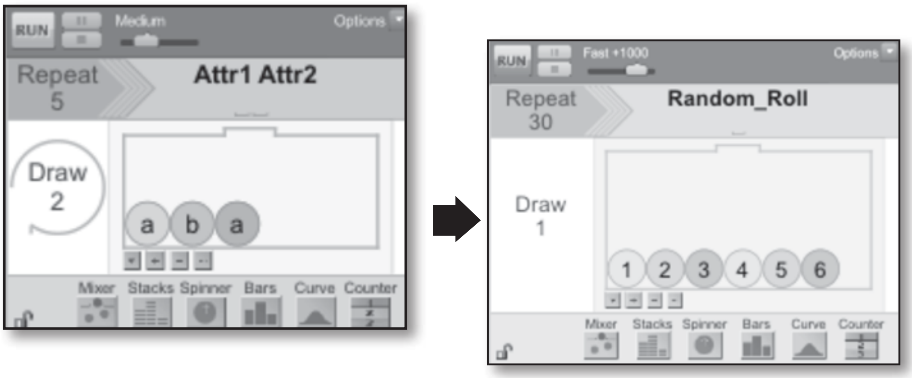
Konold, C. & Miller, C.D. (2011). *Tinkerplots: Dynamic Data Exploration* [computer software, Version 2.2]. Emeryville, CA: Key Curriculum Press. (An overview of TinkerPlots is found at http://www.srri.umass.edu/tinkerplots)

Watson, J.M. (2006). *Statistical Literacy At School: Growth And Goals.* Mahwah, NJ: Lawrence Erlbaum.

Watson, J.M., Fitzallen, N.E., Wilson, K.G. & Creed, J.F. (2008). The representational value of hats. *Mathematics Teaching in the Middle School, 14*, 4–10.

Appendix

The Sampler in *TinkerPlots* is modified to roll a single dice 30 times.



The results of running the sampler is a table of the 30 outcomes.

Results of Sampler 1 Options

	Random_Roll	<n
24	6	
25	4	
26	2	
27	4	
28	3	
29	3	
30	3	

These are dislayed in a plot.